情報量規準

の

多項式回帰モデル

①

(1) 統計的モデル

(2) 情報量規準

(3) 一般化情報量規準

GIC

ー発表内容ー

(1)定義と性質

- 最尤法と最尤推定量
- K-L情報量
- 平均対数尤度とその推定量
- 情報量規準AIC

(2)多項式回帰モデルを使った具体例

- AICで評価
- 次数選択の重要性について

$\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_q)'$ を持つ

$\{f_{\boldsymbol{\theta}}(x) ; \boldsymbol{\theta} \in \Theta\}$; 想定モデル

$g(x)$; 真のモデル

データを $\boldsymbol{x}_n = \{x_1, x_2, \cdots, x_n\}$ とすると、

対数尤度関数

$$\boxed{\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f_{\boldsymbol{\theta}}(x_i)}$$

⇒ 最尤推定量 $\hat{\boldsymbol{\theta}}$ を求める

$g(x)$；真のモデル，確率密度関数

$\{f_\theta(x)\,;\,\theta\in\Theta\}$；想定したモデル

K−L情報量

$$I(g;f_\theta)=E_g\left[\log\frac{g(x)}{f_\theta(x)}\right]$$

(i) $I(g;f_\theta)\geq 0$

(ii) $I(g;f_\theta)=0 \iff g(x)=f_\theta(x)$

$$I(g:f_{\hat{\theta}}) = \underbrace{E_g[\log g(X)]}_{\hat{\theta}\text{に依存しない}} - \underbrace{E_g[\log f_{\hat{\theta}}(X)]}_{\text{平均対数尤度}}$$

データ $x_n = \{x_1, x_2, \cdots, x_n\}$: $g(x)$ からのランダム標本

平均対数尤度の推定値

$E_g[\log f_{\hat{\theta}}(X)]$ を $\underbrace{\frac{1}{n}\sum_{\alpha=1}^{n} \log f_{\hat{\theta}}(x_\alpha)}_{\text{最大対数尤度}}$ とおく。

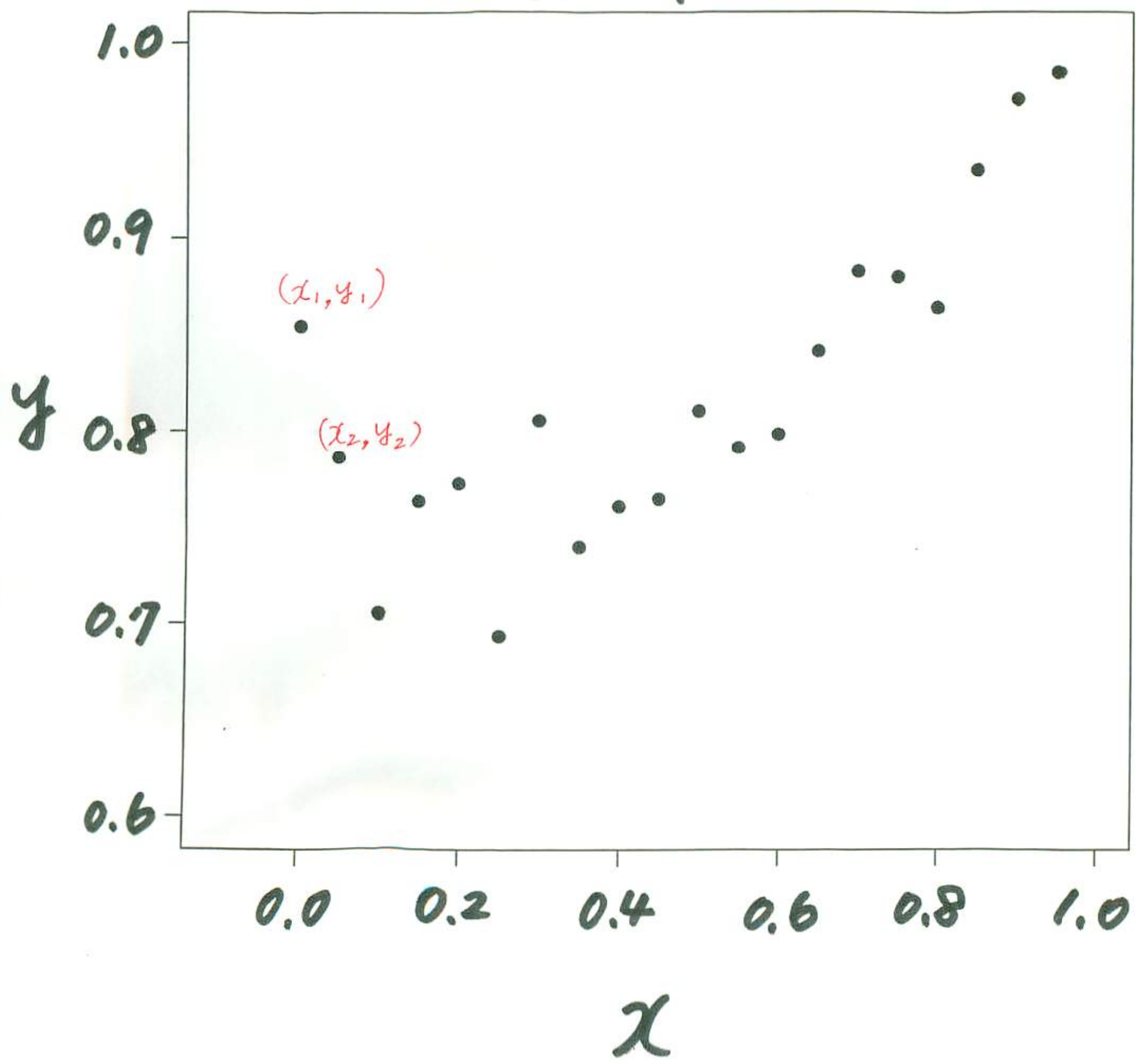
1. $$E_g\Big[\underbrace{\sum_{\alpha=1}^{n}\log f_{\hat\theta}(X_\alpha)}_{\text{最大対数尤度}} - n\underbrace{E_g \log f_{\hat\theta}(Z)}_{\text{平均対数尤度}}\Big]$$

$$\approx \underbrace{P}_{\text{バイアス}}$$ ← 漸近的に

2. $$AIC = -2\underbrace{\sum_{\alpha=1}^{n}\log f_{\hat\theta}(X_\alpha)}_{\text{最大対数尤度}} + 2\underbrace{P}_{\text{バイアス}}$$

3. AIC = −2(最大対数尤度) +2(モデルの自由パラメータ数)

散布図



j	x_j	y_j
1	0.00	0.854
2	0.05	0.786
⋮	⋮	⋮
10	0.95	0.985

多項式回帰モデル

$$y_j = \beta_0 + \beta_1 x_j + \beta_2 x_j^2 + \cdots + \beta_p x_j^p + \varepsilon_j$$

$$j = 1, \cdots, 20, \quad \varepsilon_j \sim N(0, \sigma^2)$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_{20} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{20} & x_{20}^2 & \cdots & x_{20}^p \end{pmatrix}}_{= X} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{20} \end{pmatrix}$$

最小自乗推定量は

$$\underbrace{\begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix}}_{= \hat{\beta}} = (X'X)^{-1} X' \begin{pmatrix} y_1 \\ \vdots \\ y_{20} \end{pmatrix}$$

$$AIC_p = \underbrace{20\left(\log(2\pi)+1\right)+20\log\hat{\sigma}^2}_{-2(\text{最大対数尤度})}$$

$$+2(\underbrace{P+2}_{\text{自由パラメータ数}})$$

自由パラメータ数 $<$ $\underbrace{\beta_0,\cdots,\beta_P}_{(P+1)}$, $\underbrace{\sigma^2}_{1}$

$$(P+1)+1=P+2$$

$$\hat{\sigma}^2=\frac{1}{20}\sum_{\alpha=1}^{20}\left(y_\alpha-\sum_{j=0}^{P}\hat{\beta}_j x_\alpha^j\right)^2$$

表 1: 多項式回帰モデルの推定結果

次数 p	$\hat{\sigma}^2$	AIC
1	0.002587	-56.3876
2	0.000922	-75.0268
3	0.000833	-75.0421
4	0.000737	-75.4969
5	0.000688	-74.8885
6	0.000653	-74.0041
7	0.000622	-72.8865
8	0.000607	-71.3793
9	0.000599	-69.6621

AICは $P=4$ で最小

このとき.

$$y_j = 0.835 - 1.068x_j + 3.716x_j^2 - 4.573x_j^3 + 2.14x_j^4 + \varepsilon_j$$

$$\varepsilon_j \sim N(0, 0.737\times 10^{-3})$$

(0.00, 0.854), (0.05, 0.786), (0.10, 0.706),

…, (0.90, 0.971), (0.95, 0.985)

データを1つずつ省き、19個のデータを選択

↓

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} = (X'X)^{-1}X'y$$ を推定

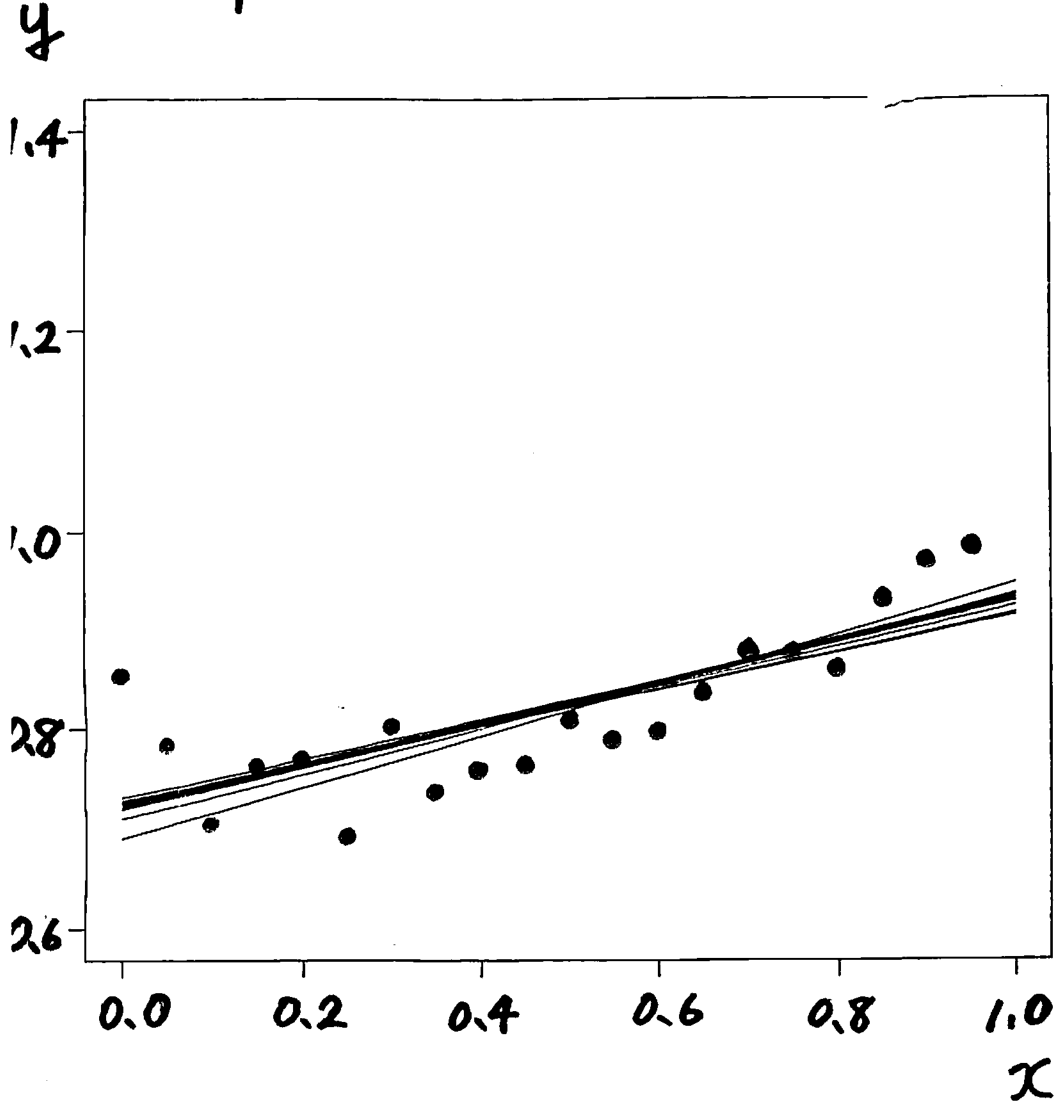
↓

$$y = \hat{\beta}_0 + \hat{\beta}_1 x + \cdots + \hat{\beta}_p x^p$$ を生成、

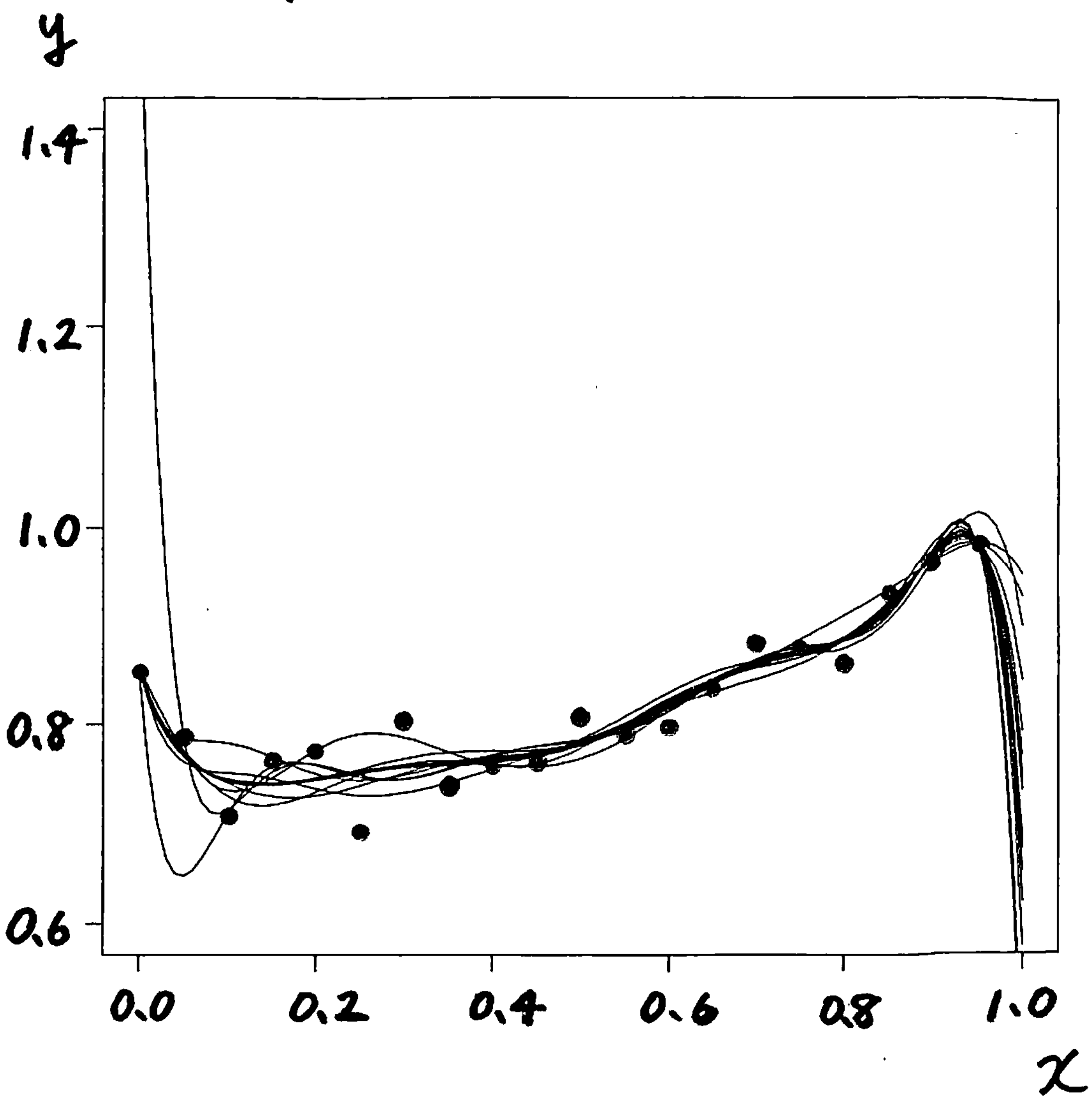
$p = 1, 4, 9$

この操作を20回繰り返す

$p=1$ のとき



$p=9$ のとき



$p = 4$ のとき